UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE
United States Patent and Trademark Office
Address: COMMISSIONER FOR PATENTS
P.O. Box 1450
Alexandria, Virginia 22313-1450
www.uspto.gov

APPLICATION NO.	FILING DATE	FIRST NAMED INVENTOR	ATTORNEY DOCKET NO.	CONFIRMATION NO.
10/626,856	07/25/2003	Thorsten H. Brants	A3052-US-NP XERZ 2 01563	9802

61962 7590 04/20/2007

FAY SHARPE / XEROX - PARC
1100 SUPERIOR AVENUE
SUITE 700
CLEVELAND, OH 44114

EXAMINER	
LOVEL, KIMBERLY M	
ART UNIT	**PAPER NUMBER**
2167	

SHORTENED STATUTORY PERIOD OF RESPONSE	MAIL DATE	DELIVERY MODE
3 MONTHS	04/20/2007	PAPER

Please find below and/or attached an Office communication concerning this application or proceeding.

If NO period for reply is specified above, the maximum statutory period will apply and will expire 6 MONTHS from the mailing date of this communication.

Office Action Summary	Application No.	Applicant(s)	
	10/626,856	BRANTS ET AL.	
	Examiner	Art Unit	
	Kimberly Lovel	2167	

-- The MAILING DATE of this communication appears on the cover sheet with the correspondence address --

Period for Reply

A SHORTENED STATUTORY PERIOD FOR REPLY IS SET TO EXPIRE 3 MONTH(S) OR THIRTY (30) DAYS, WHICHEVER IS LONGER, FROM THE MAILING DATE OF THIS COMMUNICATION.

- Extensions of time may be available under the provisions of 37 CFR 1.136(a). In no event, however, may a reply be timely filed after SIX (6) MONTHS from the mailing date of this communication.
- If NO period for reply is specified above, the maximum statutory period will apply and will expire SIX (6) MONTHS from the mailing date of this communication.
- Failure to reply within the set or extended period for reply will, by statute, cause the application to become ABANDONED (35 U.S.C. § 133). Any reply received by the Office later than three months after the mailing date of this communication, even if timely filed, may reduce any earned patent term adjustment. See 37 CFR 1.704(b).

Status

1)☐ Responsive to communication(s) filed on *22 December 2006*.

2a)☒ This action is **FINAL**. 2b)☐ This action is non-final.

3)☐ Since this application is in condition for allowance except for formal matters, prosecution as to the merits is closed in accordance with the practice under *Ex parte Quayle*, 1935 C.D. 11, 453 O.G. 213.

Disposition of Claims

4)☒ Claim(s) *1-37* is/are pending in the application.

4a) Of the above claim(s) _____ is/are withdrawn from consideration.

5)☐ Claim(s) _____ is/are allowed.

6)☒ Claim(s) *1-37* is/are rejected.

7)☐ Claim(s) _____ is/are objected to.

8)☐ Claim(s) _____ are subject to restriction and/or election requirement.

Application Papers

9)☐ The specification is objected to by the Examiner.

10)☐ The drawing(s) filed on _____ is/are: a)☐ accepted or b)☐ objected to by the Examiner.

Applicant may not request that any objection to the drawing(s) be held in abeyance. See 37 CFR 1.85(a).

Replacement drawing sheet(s) including the correction is required if the drawing(s) is objected to. See 37 CFR 1.121(d).

11)☐ The oath or declaration is objected to by the Examiner. Note the attached Office Action or form PTO-152.

Priority under 35 U.S.C. § 119

12)☐ Acknowledgment is made of a claim for foreign priority under 35 U.S.C. § 119(a)-(d) or (f).

a)☐ All b)☐ Some * c)☐ None of:

1.☐ Certified copies of the priority documents have been received.

2.☐ Certified copies of the priority documents have been received in Application No. _____.

3.☐ Copies of the certified copies of the priority documents have been received in this National Stage application from the International Bureau (PCT Rule 17.2(a)).

* See the attached detailed Office action for a list of the certified copies not received.

Attachment(s)

1) ☒ Notice of References Cited (PTO-892)
2) ☐ Notice of Draftsperson's Patent Drawing Review (PTO-948)
3) ☐ Information Disclosure Statement(s) (PTO/SB/08) Paper No(s)/Mail Date _____.
4) ☐ Interview Summary (PTO-413) Paper No(s)/Mail Date. _____ .
5) ☐ Notice of Informal Patent Application
6) ☐ Other: _____.

DETAILED ACTION

1. This communication is responsive to the Amendment filed 22 December 2006.

2. Claims 1-37 are pending in this application. Claims 1, 16, 31, 32 and 35 are independent. In the Amendment filed 22 December 2006, claims 19 and 35 have been amended. This action is made Final.

3. The rejections of claims 1-32 and 37 as being unpatentable over the article "Topic Detection and Tracking Pilot Study Final Report" by Allan et al (hereafter Final Report) in view of the article "Relevance Models for Topic Detection and Tracking" by Lavrenko et al have been withdrawn as necessitated by the arguments. The rejections of claims 33-34 as being anticipated by the article "Relevance Models for Topic Detection and Tracking" by Lavrenko et al have been withdrawn as necessitated by the amendment. The rejections of claims 35-36 being anticipated by the article have been with based on the article "Relevance Models for Topic Detection and Tracking" by Lavrenko et al have been maintained.

Claim Objections

4. The objections to claims 1 and 35 are withdrawn.

Claim Rejections - 35 USC § 102

5. The following is a quotation of the appropriate paragraphs of 35 U.S.C. 102 that form the basis for the rejections under this section made in this Office action:

> A person shall be entitled to a patent unless –
>
> (a) the invention was known or used by others in this country, or patented or described in a printed publication in this or a foreign country, before the invention thereof by the applicant for a patent.

6. Claims 35-36 are rejected under 35 U.S.C. 102(a) as being anticipated by the article "Relevance Models for Topic Detection and Tracking" by Lavrenko et al (hereafter Relevance Model).

Referring to claim 35, Relevance Model teaches a computer-implemented method of detecting new events comprising the steps of:

determining a first source-identified story [story S1] associated with at least one event (Relevance Model: see section 3.2: Measuring Topic Similarity);

determining a second source-identified story associated with at least one event (Relevance Model: see section 3.2: Measuring Topic Similarity – story S2);

determining a story-pair based on the first source-identified story and the second source-identified story (Relevance Model: see section 3.2: Measuring Topic Similarity, lines 1-14 – S1 and S2);

outputting an indicator of inter-story similarity between the first and second story based on at least one of:

an event frequency model, story segmentation (Relevance Model: see section 1: Introduction, lines 3-7) and a source-identified inter-story similarity metric.

Referring to claim 36, Relevance Model discloses the method of claim 35, wherein story segmentation is based on at least one of: topic (see section 1: Introduction, lines 3-7), an adjacent window and an overlapping window.

Claim Rejections - 35 USC § 103

7. The following is a quotation of 35 U.S.C. 103(a) which forms the basis for all obviousness rejections set forth in this Office action:

> (a) A patent may not be obtained though the invention is not identically disclosed or described as set forth in section 102 of this title, if the differences between the subject matter sought to be patented and the prior art are such that the subject matter as a whole would have been obvious at the time the invention was made to a person having ordinary skill in the art to which said subject matter pertains. Patentability shall not be negatived by the manner in which the invention was made.

This application currently names joint inventors. In considering patentability of the claims under 35 U.S.C. 103(a), the examiner presumes that the subject matter of the various claims was commonly owned at the time any inventions covered therein were made absent any evidence to the contrary. Applicant is advised of the obligation under 37 CFR 1.56 to point out the inventor and invention dates of each claim that was not commonly owned at the time a later invention was made in order for the examiner to consider the applicability of 35 U.S.C. 103(c) and potential 35 U.S.C. 102(e), (f) or (g) prior art under 35 U.S.C. 103(a).

8. **Claims 1-2, 7-17, 22-32 and 37 are rejected under 35 U.S.C. 103(a) as being unpatentable over the article "Topic Detection and Tracking Pilot Study Final Report" by Allan et al (hereafter Final Report) in view of the article "Relevance Models for Topic Detection and Tracking" by Lavrenko et al.**

Referring to claim 1, Final Report discloses a method of detecting new events (see section 1.1: Background, lines 1-6) comprising the steps of:

determining at least one story characteristic based on at least one of:

an average story similarity story (see section 3.1: Detection Evaluation: lines 34-44) and a same event-same source story characteristic;

determining a source-identified story corpus, each story associated with at least one event (see section 1.2: The Corpus; and section 2.1: Evaluation, lines 11-16 – the corpus is the TDT study corpus (a source) has been developed by the TDT study effort and comprises of 16,000 stories from Reuters newswire and CNN broadcast news transcripts);

determining a source-identified new story associated with at least one event (see section 3: New Event Detection, lines 19-24 – a Yes/No decision per story made at the time when the story arrives regarding whether or not the story is the first reference to a newly reported event);

determining story-pairs based on the source-identified new-story and each story in the source-identified story corpus (see section 3.1: Detection Evaluation, lines 6-48 – each new story is associated with a cluster which is considered to represent a story-pair);

determining at least one inter-story similarity metric [score] for the story-pairs (see section 3.1: Detection Evaluation, lines 30-33 – the score indicates how confident the system is that the story being processed discusses the cluster event); and

outputting a new story event indicator [YES or NO decision] (see section 3.1: Detection Evaluation, lines 23-29) if the event associated with the new story is similar to the events associated with the source-identified story corpus based on the inter-story similarity metrics (see section 3.1: Detection Evaluation, lines 30-33 - score).

However, Final Report fails to explicitly teach the limitation of determining at least one adjustment to the inter-story similarity metrics based on at least one story characteristic and then using the adjustment to output a new story indicator. Relevance Model also discloses a method for detecting new events similar to that of Final Report including the further limitation. In particular, Relevance Model discloses determining at least one adjustment to the inter-story similarity metrics based on at least one story characteristic (see section 4.3: Relevance Model Performance – the KL metric is adjusted by a clarity value) and using the adjustment to output a new event story indicator (see section 3.2: Measuring Topic Similarity).

It would have been obvious to one of ordinary skill at the time the invention was made to utilize Relevance Model's method of determining at least one adjustment to the inter-story similarity metrics based on at least one story characteristic with Final Report's method of detecting new events by using the KL metric and the score metric. One would have been motivated to do so since the Kullback-Leibler metric provides

relevance modeling, which enhances the topic model estimate associated with a news story (Relevance Model: see abstract, lines 3-6).

Referring to claim 2, the combination of Final Report and Relevance Model (hereafter FinalReport/RelevanceModel) teaches the method of claim 1, wherein the inter-story similarity metric is adjusted based on at least one of subtraction (Relevance Model: see section 4.3: Relevance Model Performance, lines 1-13 – the symmetric clarity-adjusted KL is the similarity metric utilized; and section 4.2: Value of Clarity-adjusted KL, line 12 – the equation uses subtraction) and division (Relevance Model: see section 3.2: Measuring Topic Similarity, lines 31-37 – clarity is considered to represent the adjustment).

Referring to claim 7, FinalReport/RelevanceModel teaches the method of claim 1, wherein the inter-story similarity metrics are comprised of: at least one story frequency model (Final Report: see section 4.1: Detection Experiments, lines 13-15); and

at least one event frequency model combined using terms weights (Relevance/Model: see section 5.1: Tracking algorithm, lines 1-23).

Referring to claim 8, FinalReport/RelevanceModel teaches the method of claim 1, wherein the inter-story similarity metrics are comprised of at least one story frequency model (Final Report: see section 4.1: Detection Experiments, lines 13-15); and

at least one story characteristic frequency model combined using terms weights (Relevance Model: see Figure 6).

Referring to claim 9, FinalReport/RelevanceModel teaches the method of claim 8, where the adjustments based on the story characteristics are applied to the term weights (Relevance Model: see section 1: Introduction, lines 29-31 and Figure 6).

Referring to claim 10, FinalReport/RelevanceModel teaches the method of claim 8, where the adjustments based on the story characteristics are applied to the inter-story similarity metrics (Relevance Model: see section 4.3: Relevance Model Performance – the KL metric is adjusted by a clarity value).

Referring to claim 11, FinalReport/RelevanceModel teaches the method of claim 1, wherein the inter-story similarity metrics are comprised of at least one term frequency-inverse event frequency model (Final Report: see section 4.1: Event Detection, lines 11-23) and where the events are classified based on at least one of: story labels and a predictive model (Relevance Model: see section 5.1: Tracking Algorithm, lines 11-23).

Referring to claim 12, FinalReport/RelevanceModel teaches the method of claim 8, wherein an event frequency is determined based on term *t* and ROI category *rmax* from the formula: $ef_{r\max}(t) = \max_{r \in R}(ef(r,t))$ (Final Report: see section 3.2: Measuring Topic Similarity – the equation finds the probability of the topic).

Referring to claim 13, FinalReport/RelevanceModel teaches the method of claim 8, wherein an the inverse event frequency is determined based on term *t*, and events *e* and *rmax* in the set of ROI categories from the formula: $IEF(t) = \log\left[\frac{N_{e,r}}{ef(r,t)}\right]$

(Final Report: see section 3.5 Results, Analysis, and Future Work – finding the inverse document frequency is comparable to finding the inverse event frequency).

Referring to claim 14, FinalReport/RelevanceModel teaches the method of claim 8, wherein an inverse event frequency is determined based on term t, categories e, r and rmax in the set of ROI categories and P(r), the probability of ROI r from the formula: $IEF'(t) = \sum_{r \in R} P(r) \log\left[\frac{N_{e,r}}{ef(r,t)}\right]$ (Final Report: see section 3.5 Results, Analysis, and Future Work – finding the inverse document frequency is comparable to finding the inverse event frequency; the derivative of the first equation has been taken).

Referring to claim 15, FinalReport/RelevanceModel teaches the method of claim 1 further comprising the step of determining a subset of stories from the source-identified story corpus and the source-identified new story based on at least one story characteristic (Final Report: see section 3.1: Detection Evaluation – the stories within the corpus are placed in clusters which is considered to represent a subset).

Referring to claim 16, the claim is rejected on the same grounds as claim 1.

Referring to claim 17, the claim is rejected on the same grounds as claim 2.

Referring to claim 22, the claim is rejected on the same grounds as claim 7.

Referring to claim 23, the claim is rejected on the same grounds as claim 8.

Referring to claim 24, the claim is rejected on the same grounds as claim 9.

Referring to claim 25, the claim is rejected on the same grounds as claim 10.

Referring to claim 26, the claim is rejected on the same grounds as claim 11.

Referring to claim 27, the claim is rejected on the same grounds as claim 12.

Referring to claim 28, the claim is rejected on the same grounds as claim 13.

Referring to claim 29, the claim is rejected on the same grounds as claim 14.

Referring to claim 30, the claim is rejected on the same grounds as claim 15.

Referring to claim 31, the claim is rejected on the same grounds as claim 1.

Referring to claim 32, the claim is rejected on the same grounds as claim 1.

Referring to claim 33, FinalReport/RelevanceModel discloses the method of claim 1, inter-story similarity information combined comprising the steps of:

determining P(sameROI(q,d)) based on the probability of story q and story d having the same ROI category (Relevance Model: see section 3.2: Measuring Topic Similarity – according to paragraph [0013], lines 6-7 of the applicant's specification, the definition of an ROI category is a categorization of events; the topic is considered to represent an event);

determining $similarity_{IEF''}$, based on a similarity with no inverse event frequency influence (Relevance Model: see section 3.2: Measuring Topic Similarity); and

the formula: $similarity'(q,d)=P(sameROI(q,d))* similarity_{IEF'}(q,d)+(1-P(sameROI(q,d)))* similarity_{IEF''}(q,d)$ (Relevance Model: see section 3.2: Measuring Topic Similarity, lines 15-17 – the formula is an extended version of D(M1||M2) + D(M2||M1)).

Referring to claim 34, FinalReport/RelevanceModel teaches the method of claim 33, wherein P(sameROI(q,d)) is based on the formula:

$$P(sameROI(q,d)) = \frac{N_{same}(similarity_{IEF''}(q,d))}{N_{same}(similarity_{IEF''}(q,d)) + N_{different}((similarity_{IEF''}(q,d))}$$ (Relevance

Model: see section 3.2: Measuring Topic Similarity, lines 31-44 – clarity represents Ndiff).

Referring to claim 37, FinalReport/RelevanceModel discloses the computer-implemented method of claim 1, in which the new event indicator is displayed on at least one of a visual, audio or tactile output device (Final Report: Section 3.6: Open Issues).

9. Claims 3-5 and 18-20 are rejected under 35 U.S.C. 103(a) as being unpatentable over the article "Topic Detection and Tracking Pilot Study Final Report" by Allan et al in view of the article "Relevance Models for Topic Detection and Tracking" by Lavrenko et al as applied respectively to claims 1 and 16 above, and further in view of US Patent No 6,584,220 issued to Lantrip et al (hereafter Lantrip et al).

Referring to claim 3, FinalReport/RelevanceModel teaches a method for new event detection as claimed in claim 1. FinalReport/RelevanceModel also teaches the further limitation wherein the inter-story similarity metric is at least one of a probability based inter-story similarity metric (see section 3.2: Measuring Topic Similarity, lines 1-14). However, FinalReport/RelevanceModel does not explicitly teach the further limitation wherein the inter-story similarity metric is also a Euclidean based inter-story similarity metric. Lantrip et al teaches a method of similarity metrics including a Euclidean based inter-story similarity metric (Lantrip et al: see column 4, lines 46-48).

It would have been obvious to one of ordinary skill at the time the invention was made to utilize Lantrip et al's method of a Euclidean based inter-story metric with Final Report/RelevanceModel's method of a probability based inter-story metric. One would have been motivated to do so since a Euclidean based inter-story metric aids in solving the problem of how to train and cluster a corpus (Final Report: see section 1.2).

Referring to claim 4, the combination of (FinalReport/RelevanceModel and Lantrip et al (hereafter FinalReport/RelevanceModel/Lantrip) teaches the method of claim 3, wherein the probability based inter-story similarity metric is at least one of a Hellinger, a Tanimoto, a KL divergence (Relevance Models: see section 3.2: Measuring Topic Similarity, lines 1-14 – Kullback-Leibler divergence) and a clarity distance based metric.

Referring to claim 5, FinalReport/RelevanceModel/Lantrip teaches the method of claim 3, wherein the Euclidean based similarity metric is a cosine-distance based metric (Lantrip et al: see column 4, lines 46-48).

Referring to claim 18, the claim is rejected on the same grounds as claim 3.

Referring to claim 19, the claim is rejected on the same grounds as claim 4.

Referring to claim 20, the claim is rejected on the same grounds as claim 5.

10. Claims 6 and 21 are rejected under 35 U.S.C. 103(a) as being unpatentable over the article "Topic Detection and Tracking Pilot Study Final Report" by Allan et al in view of the article "Relevance Models for Topic Detection and Tracking" by Lavrenko et al as applied respectively to claims 1 and 16 above, and further in view of the article "On-line New Event Detection and Tracking" by Allan et al (hereafter New Event Detection).

Referring to claim 6, FinalReport/RelevanceModel teaches a method of detecting new events. However FinalReport/RelevanceModel fails to explicitly teach the further limitation wherein the inter-story metrics are determined based on a term frequency-inverse story frequency model. New Event Detection discloses a method similar to that of FinalReport/RelevanceModel. New Event Detection discloses a method similar to that of FinalReport/RelevanceModel including the further limitation.

In particular, New Event Detection discloses a method similar to that of claim 1, wherein the inter-story similarity metrics are determined based on a term frequency-inverse story frequency model (see section 4.1: Detection Experiments, lines 13-15).

It would have been obvious to one of ordinary skill at the time the invention was made to utilize New Event Detection's method of a term frequency-inverse story frequency with Final Report/RelevanceModel's method of an inter-story similarity metrics. One would have been motivated to do so since all three articles discuss the TDT initiative carried out by the Center for Intelligent Information Retrieval (Final Report: see abstract; Relevance Model: see abstract; New Event Detection: see abstract).

Referring to claim 21, the claim is rejected on the same grounds as claim 6.

Response to Arguments

11. Applicant's arguments filed 22 December 2006 have been fully considered but they are not persuasive.

12. In regards to applicant's arguments on pages 10-11 concerning the prior art rejection of claims 35-36, applicant states: There is simply no discussion in Lavrenko of source-identified sources nor of a source identified inter-story similarity metric.

The examiner respectfully disagrees.

The limitations in the claim state a first source-identified story and a second source-identified story and not source-identified sources. Lavrenko discloses a TDT2 dataset which contains 40,000 news stories which were collected from six different sources: two newswire sources, two radio sources and two television sources. The term "source-identified story" is not explicitly defined in the specification. Therefore, when interpretating the term in the broadest sense, Lavrenko discloses source-identified stories in section 4.1.1: Dataset since Lavrenko discloses the sources of the stories located within the corpus.

Lavrenko discloses a inter-story similarity metric in section 3.2: Measuring Topic Similarity. Lavrenko uses Kullback-Leibler divergence to determine that the models of the first story and the second story represent the same topic. The relevance model for each story is intended to capture the topic of the stories (see section 3: Relevance Models in TDT). Therefore, using the models to determine the similarity between two stories is considered to represent an "an indicator of inter-story similarity between the first and second story."

13. In regards to applicant's arguments on pages 12-13 concerning the prior art rejection of claims 1-2, 7-17, 22-34 and 37, applicant states: The communication at p 5, lines 11-13 asserts that the features in the claims directed to: 1) determining at least one story characteristic based on at least one of: an average story similarity story characteristic and a same event-same source story characteristic are descried in Allan1, Section 3.1, lines 334-44.

The citation is meant to be directed towards the limitation of an average story similarity story characteristic. Each story is compared to all of the stories located in the cluster. Furthermore, Allan1 does teach the concept of source-identified stories in section 2.1: The Corpus. Allan1 states that the corpus consists of 16,000 stories with half from Rueters newswire and half from CNN broadcast. When interpretating the term "source-identified story corpus" in the broadest sense, Allan1's story corpus can be utilized to represent a source-identified story corpus.

Conclusion

14. **THIS ACTION IS MADE FINAL.** Applicant is reminded of the extension of time policy as set forth in 37 CFR 1.136(a).

A shortened statutory period for reply to this final action is set to expire THREE MONTHS from the mailing date of this action. In the event a first reply is filed within TWO MONTHS of the mailing date of this final action and the advisory action is not mailed until after the end of the THREE-MONTH shortened statutory period, then the shortened statutory period will expire on the date the advisory action is mailed, and any extension fee pursuant to 37 CFR 1.136(a) will be calculated from the mailing date of the advisory action. In no event, however, will the statutory period for reply expire later than SIX MONTHS from the mailing date of this final action.

Contact Information

Any inquiry concerning this communication or earlier communications from the examiner should be directed to Kimberly Lovel whose telephone number is (571) 272-2750. The examiner can normally be reached on 8:00 - 4:00.

If attempts to reach the examiner by telephone are unsuccessful, the examiner's supervisor, John Cottingham can be reached on (571) 272-7079. The fax phone number for the organization where this application or proceeding is assigned is 571-273-8300.

Information regarding the status of an application may be obtained from the Patent Application Information Retrieval (PAIR) system. Status information for published applications may be obtained from either Private PAIR or Public PAIR. Status information for unpublished applications is available through Private PAIR only. For more information about the PAIR system, see http://pair-direct.uspto.gov. Should you have questions on access to the Private PAIR system, contact the Electronic Business Center (EBC) at 866-217-9197 (toll-free). If you would like assistance from a USPTO Customer Service Representative or access to the automated information system, call 800-786-9199 (IN USA OR CANADA) or 571-272-1000.

Kimberly Lovel
Examiner
Art Unit 2167

13 April 2007
kml

JOHN COTTINGHAM
SUPERVISORY PATENT EXAMINER
TECHNOLOGY CENTER 2100